INSTITUTE for REPLICATION

No. 9
I4R DISCUSSION PAPER SERIES

Counting Missing Women – A Reconciliation of the 'Flow Measure' and the 'Stock Measure'

Cara Ebert

Stephan Klasen

Sebastian Vollmer

November 2022

I4R Discussion Paper Series

I4R DP No. 9

Counting Missing Women – A Reconciliation of the 'Flow Measure' and the 'Stock Measure'

Cara Ebert[1], Stephan Klasen[2], Sebastian Vollmer[2]

[1]RWI – Leibniz Institute for Economic Research, Berlin/Germany
[2]University of Göttingen/Germany

November 2022

E-Mail: joerg.peters@rwi-essen.de
RWI – Leibniz Institute for Economic Research
Hohenzollernstraße 1-3
45128 Essen/Germany
www.i4replication.org

ISSN: 2752-1931

Counting Missing Women – A Reconciliation of the `Flow Measure' and the `Stock Measure'

Cara Ebert
RWI – Leibniz Institute for Economic Research

Stephan Klasen[1]
University of Göttingen

Sebastian Vollmer[2]
University of Göttingen

Abstract

`Stock estimates' of missing women suggest that the problem is concentrated in South and East Asia and among young children. In contrast, `flow estimates' suggest that gender bias in mortality is much larger, is as severe among adults as it is among children in India and China, and is larger in Sub-Saharan Africa than in India and China. We show that the different stock and flow measure results rely on the choice of the reference standard for mortality and an incomplete correction for different disease environments in the flow measure. Alternative reference standards reconcile the results of the two measures.

JEL Codes: J16, D63, I10
Keywords: Missing women, gender bias, mortality, disease, age, Sub-Saharan Africa, China, India

Acknowledgements: We would like to thank Jere Behrman, Sonia Bhalotra, Monica Das Gupta, Angus Deaton, Ken Harttgen, Alan Lopez, Ashok Rai, Debraj Ray as well as participants at seminars and workshops at Harvard, UPenn, Tufts, Dalhousie, and Göttingen for helpful comments and discussion.

[1] Stephan Klasen passed away on October 27, 2020 after battling the incurable disease Amyotrophic Lateral Sclerosis (ALS) for 5 years. We are grateful for what we have learned from Stephan through our joint work, and previously as students.

[2] Corresponding author: University of Göttingen, Waldweg 26, 37073 Göttingen, Germany.
Email: svollmer@uni-goettingen.de.

1. Introduction

The extent of gender bias in mortality in developing countries has been a research topic that sparked substantial controversies in the economic and demographic literature since the time when Amartya Sen first pointed to the millions of `missing women' in South and East Asia in 1989 (Sen, 1989). This literature focused largely on calculating stock measures of missing women. Stock measures compare the actual sex ratio (males/ females), using census information in a country, with an expected sex ratio that would occur in the absence of gender bias in mortality. An alternative method to the stock measure poses the flow measure of missing women (Anderson and Ray, 2010). It compares the ratio of male to female age-specific mortality rates with an expected ratio of male to female age-specific mortality rates from a reference population without gender bias. The flow measure estimates excess female mortality in a given year, whereas the stock measure estimates the total deficit in alive women at one point in time. These two measures have resulted in very different implications about the overall number of missing women and their distribution by age and region. This paper examines the flow measure's sensitivity to alternative assumptions and methods, and reconciles the findings from the two measures.

Missing women estimates based on the stock measure, and many more studies on gender inequalities using micro data from individual countries or regions, have converged on two key findings (e.g. Sen, 1989, Coale, 1991; Bannister and Coale, 1994; Das Gupta, 1987, Murthi, Guio and Dreze, 1995; Das Gupta, 2005; Klasen and Wink, 2002, 2003; Das Gupta, Chung, and Shuzhou, 2009). First, the problem of missing women is much more prevalent in China and India than in Sub-Saharan Africa. Second, in China and India, missing women are largely attributable to excess female mortality before birth and/or the first few years of life. The stock estimates amounted to around 100 million missing women in 2000 in the most affected countries, i.e. China, India, Pakistan, Bangladesh, the countries of Western Asia and some countries of North Africa (Klasen and Wink, 2003). Compared with a previous census round in 1990, the number of missing women had increased by about 6 million between 1990 and 2000 but fallen as a share of females alive (from 6.5% to 5.7%) (Klasen and Wink, 2003). [1]

Contrary to findings from the stock measure, estimates of the annual flow of missing women for the year 2000 in China, India, and Sub-Saharan Africa suggest that excess female mortality is (a) as severe among adults as it is among children in India and China, (b) higher in Sub-Saharan Africa than in China and India and (c) largely due to excess adult female mortality in Sub-Saharan Africa (Anderson and Ray, 2010; World Bank, 2012). For China, India and Sub-Saharan Africa together, excess mortality based on the flow measure is also much larger than implied by the existing literature on stock measures. The flow measure suggests that in a single year about four to five million women died in excess in India, China, and Sub-Saharan Africa alone. For the 1990s, the stock measure would be consistent with a flow of about one million missing women in a given year.[2] Further, flow measure estimates show that excess female mortality in developing countries is elevated among unmarried women, particularly in the reproductive age (15-49) (Anderson and Ray, 2019).

This paper discusses three key methodological aspects of the flow measure and shows that the flow measure results are consistent with the stock measure results when a set of different assumptions are made. The first two points regard the reference standard, which is used to simulate a gender equal society, and the third point regards the availability of quality mortality data required for the calculation of the flow of missing women. The ideal counterfactual for assessing the amount of gender bias in mortality would be a society in which males and females are treated equally and in which there are no differences in sex-

[1] A major controversy about these stock estimates of missing females arose when Oster (2005) claimed that the 'natural' sex ratio at birth in some of these countries was much higher due to the fact that the widespread prevalence of the hepatitis-B-carrier status raised the sex ratio at birth. She claimed that this way some 50-70% of missing females could be accounted for, i.e. their absence was unrelated to discrimination. After an intense debate and further research (Abrevaya, 2009; Das Gupta, 2005, 2006; Lin and Luoh, 2008), Oster retracted the finding in 2009 (Chen and Oster 2010) and the previous estimates of around 100 million were re-instated. See Klasen (2008) for a review of this debate.

[2] Note that flows of missing cannot be calculated as differences in stocks between time periods, i.e. the difference of 6 million in stocks does not imply 600,000 missing women per year. This is due to the fact that some women in a particular age group that have died in excess in a year might not count as missing in a stock measure at the next census if in the time interval some men in that age group have also died due to high overall mortality, i.e. to the extent that an excess female death in a year just advances mortality between two census periods, it would not add to the increase in the stock of missing women at the next census.

based behavioral patterns across countries with significant mortality implications (Coale, 1991, Klasen and Wink, 2002, 2003, Anderson and Ray, 2010). There is no such society past or present that exhibits this feature so that it is inherently difficult to generate a reference standard that is entirely beyond reproach. However, because it is so difficult to identify an adequate reference standard, it is important to understand the implications of the choices made. Throughout the paper, we refer to the flow measure based on the methods and data used in Anderson and Ray (2010) – the seminal paper on the measurement of the annual flow of missing women – and therefore the number of missing women in 2000.

First, we review the reference population of high-income countries in 2000 used in flow estimates.[3] Today's high-income countries are an environment with extremely low overall adult mortality and, therefore, already small differences in male and female mortality rates result in large changes in the male to female mortality rate of high-income countries. This is empirically relevant as non-natural causes of death or consequences of behavioral patterns such as accidents, violence, suicides and smoking affect males much more than females. Second, we discuss that disease environments do not differ across countries for males and females in the same way, but they differ across sexes and countries. Therefore, controlling for differences in the disease environment in the reference standard requires to control for sex differences within disease groups across countries. For example, the male to female mortality ratio from AIDS is around 4 in high-income countries in 2000 because AIDS mortality was confined to male-dominated high-risk groups (gay men and IV drug users), whereas in Sub-Saharan Africa AIDS spread from the beginning to the entire population, with heterosexual transmission playing a key role, and AIDS mortality being more similar for males and females. Thus, using today's high-income countries as a reference population for AIDS mortality suggests massive excess female mortality in adulthood in countries with high rates of AIDS mortality. The World Development Report 2012, which uses the flow measure approach, indeed shows particularly high rates and increases in excess female mortality in AIDS affected countries of Sub-Saharan Africa (World Bank, 2012). The two points on the low mortality environment and the disease correction are related, but each have their separate effect on the estimates of missing

[3] High-income countries refer to the former World Bank country classification 'Established Market Economies' as used by Anderson and Ray (2010), which includes Western Europe, Canada, United States, Australia, New Zealand, and Japan.

women and are therefore also discussed separately. In a third point, we highlight a general concern that underlies the use of mortality data in Sub-Saharan Africa which is that reliable vital registration data is only available in one of 46 countries. For the remaining countries, mortality data must be estimated and, therefore, the flow of missing women is inherently imprecise.

We use a variety of alternative reference standards for sex-specific mortality and find patterns in the flow of missing women for the year 2000 that are consistent with previous findings on the stock of missing women. We find that excess female mortality is much more serious in India and China than in Sub-Saharan Africa. In China and India, excess female mortality is largely driven by gender bias pre-birth and during the first few years of life. The magnitude of the problem is also much reduced, our estimates suggest that 1.5-1.8 million women die in excess in the three regions in 2000. This number is substantial and worrying but below the number of five million females per year estimated in Anderson and Ray (2010) or 3.3 million below the age of 60 of the World Bank (2012). Our estimate is also lower, but to a lesser extent, than the flow measure estimate by Bongaarts and Guilmoto (2015), who show trends in excess female mortality from 1970 to 2050 using data from 93 countries with different mortality levels to create a reference population.[4] Bongaarts and Guilmoto (2015) do not provide flow estimates by age such that comparisons of age patterns cannot be made. Further, our calculations confirm severe gender bias in mortality among adults in Sub-Saharan Africa, likely linked to elevated mortality of women due to AIDS, although it is of smaller magnitude than the original flow estimates.

[4] The estimations of Bongaarts and Guilmoto (2015) amount to excess female deaths after birth of 0.54 million in China and 0.82 million in India in 2000. They predict the reference sex ratio of mortality based on a linear regression of the log sex ratio on life expectancy for combinations of different age groups and years (304 separate regressions). The countries included in the computation of the reference population are Algeria, Angola, Argentina, Australia, Austria, Belgium, Benin, Bolivia, Brazil, Burkina Faso, Burundi, Cambodia, Cameroon, Canada, Chad, Chile, Colombia, Côte d'Ivoire, Cuba, Democratic Republic of the Congo, Denmark, Dominican Republic, Ecuador, Egypt, El Salvador, Eritrea, Ethiopia, Finland, France, Germany, Ghana, Greece, Guatemala, Guinea, Haiti, Honduras, Indonesia, Iran, Iraq, Israel, Italy, Japan, Jordan, Kazakhstan, Kenya, Kyrgyzstan, Laos, Libya, Madagascar, Malawi, Mali, Mexico, Morocco, Mozambique, Myanmar, Netherlands, Nicaragua, Niger, Nigeria, North Korea, Papua New Guinea, Paraguay, Peru, Philippines, Portugal, Rwanda, Senegal, Serbia, Sierra Leone, Somalia, South Africa, South Sudan, Spain, Sri Lanka, Sudan, Sweden, Switzerland, Syria, Tajikistan, Thailand, Togo, Tunisia, Turkey, Turkmenistan, Uganda, United Kingdom, Tanzania, United States, Uzbekistan, Venezuela, Yemen, Zambia, and Zimbabwe.

Our results highlight the sensitivity of missing women estimates to alternative assumptions and methods, and suggest that the problem of missing women, as previously found, likely still is largest among children and infants in India and China. This is important because missing women estimates guide policy-making. As evident by the World Development Report on 'Gender Equality and Development', the flow of missing women poses a relevant metric to highlight grievances in gender equality in developing countries. Flow estimates allow for the identification of specific age groups at risk of excess mortality and direct comparisons of missing women over time as well as attributions of deaths to specific disease groups. The World Development Report concludes “Excess female mortality is slowly shifting from early childhood in South Asia to adulthood in Sub-Saharan Africa, declining in all low-income countries except in Sub-Saharan Africa” (World Bank, 2012, p.77). The report also points out that some of the results were not detected or are even in contrast with earlier stock measure estimates: “Less well known is that excess female mortality is a continuing phenomenon beyond childhood and a growing problem in Sub-Saharan Africa. […] While missing girls at birth are indeed concentrated in India and China, consistent with the earlier discussion, excess female mortality after birth is highest in Sub-Saharan Africa, the only region where the numbers are going up over time” (World Bank, 2012, pp.118-120). While our flow estimates differ to previous ones in the size of missing women aged 15 to 25 in Sub-Sahara Africa, they do reinforce that the number of young missing women is large and should be subject of future research.

This paper is organized as follows. In section 2, we present the conceptual framework of the flow measure of missing women. In section 3, we discuss the problem of low mortality reference populations and the sex ratio as a functional form of the reference standard, and present estimates of missing women using alternative reference populations and functional forms. In section 4, we discuss the relevance of sex-specific differences within diseases across countries and present estimates of missing women using an alternative reference standard for AIDS. In section 5, we draw attention to the issue of lacking mortality data in Sub-Saharan Africa. Section 6 concludes.

2. The 'flow measure' of missing women

Using the flow measure, excess female mortality (EFM) for age group a can be calculated the following way:

$$EFM(a) = \left(MR_f(a) - \left(\frac{MR_m(a)}{\widehat{MR_m}(a) / \widehat{MR_f}(a)} \right) \right) * Pop_f(a) \qquad (1)$$

$MR_f(a)$ and $MR_m(a)$ refer to the *actual* mortality rates of females and males in age group (a), respectively. $\widehat{MR_f}(a)$ and $\widehat{MR_m}(a)$ describe the mortality rates of females and males in age group (a) in the *reference population*, respectively. $\widehat{MR_m}(a)/\widehat{MR_f}(a)$ is the male to female sex ratio of the reference population and is used as a correction factor to scale the actual male mortality rate $MR_m(a)$ to account for, for example, differences in mortality from risky behavior or specific diseases across sexes. The scaled $MR_m(a)$ provides the *reference mortality rate of females* in age group a which is subtracted from *the actual mortality rate of females* to receive the population share of missing women in age group a. Multiplying the population share of missing women in age group a with the number of women in age group a, $Pop_f(a)$, results in the total flow of missing women in age group a, $EFM(a)$. Based on equation (1), the essential ingredients to calculate excess female mortality are the age-specific actual mortality rates and the correction factor consisting of the age-specific male to female mortality rate ratio of the reference population. As a reference population, the flow measure of missing women uses high-income countries in 2000. The reference population and the correction method together build the reference standard.

Table 1 presents the mortality rates for adults aged 15 to 59 by sex for Sub-Saharan Africa, India and China in 2000 as actual mortality rates ($MR_f(15-59)$ and $MR_m(15-59)$) and for high-income countries in 2000 as reference mortality rates ($\widehat{MR_m}(15-59)$ and $\widehat{MR_f}(15-59)$). The actual sex ratios of mortality rates range between 1.13 and 1.46. Adult males thus die at 13-46% higher rates than adult females. The sex ratio in high-income countries in 2000 is about 1.92. Thus, adult males die at 92% higher rates than adult females. This sex ratio in high-income countries constitutes the key component of the flow measure's reference standard and the estimates of missing women will be highly sensitive to this reference standard.

Table 1: Adult mortality rates (15-59) by sex in 2000

	(1) Male	(2) Female	(3) Sex Ratio
Sub-Saharan Africa	0.511	0.454	1.125
India	0.287	0.213	1.347
China	0.161	0.110	1.464
High-income countries	0.125	0.065	1.918

Notes: Adult mortality rates refer to the likelihood of having died by age 60 if one was alive at age 15. Male refers to the adult (15-59) mortality rate among males. Female refers to the adult (15-59) mortality rate among females. Sex ratio refers to the ratio of columns (1) and (2), i.e. the male to female mortality rate ratios. To derive mortality rates in Sub-Saharan Africa and high-income countries, population (15-59 years old) weighted averages of country specific adult mortality rates were calculated (see supplementary materials). Sub-Saharan African countries include: Algeria, Angola, Benin, Botswana, Burkina Faso, Burundi, Cameroon, Cape Verde, CAR Chad, Comoros, Cote d'Ivoire, DRC Equatorial Guinea, Eritrea, Ethiopia, Gabon, Gambia, Guinea, Guinea-Bissau, Kenya, Lesotho, Liberia, Madagascar, Malawi, Mali, Mauritania, Mauritius, Mozambique, Namibia, Niger, Nigeria, Rwanda, Sao Tome and Principe, Senegal, Seychelles, Sierra Leone, South Africa, Swaziland, Tanzania, Togo, Uganda, Zambia, Zimbabwe. High-income countries include: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Iceland, Ireland, Italy, Japan, Luxembourg, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom, United States.
Sources: Country specific adult mortality rates in 2000 were taken from WHR (2001). Population numbers were taken from The UN World Population Prospect Revision 2000: https://population.un.org/wpp/Download/Archive/Standard/, downloaded April 3, 2021.

Following equation (1) and using high-income countries in 2000 as a reference population, Table 2 presents the flow of missing women in Sub-Saharan Africa, India and China in 2000 (Anderson and Ray, 2010).[5] With this reference standard, the problem of missing women as a share of the female population is worst in Sub-Saharan Africa. 0.47% of females died in excess in 2000 in Sub-Saharan Africa, compared to 0.34% in India and 0.31% in China. Table 2 further illustrates that in India and China there is a massive problem of missing women among adults that is larger than among children. At last, adding the total number of missing women across Sub-Saharan Africa, India and China results in 5 million females that went missing in one year.

[5] For example, in Sub-Saharan Africa, the mortality rate for females aged 1-5 is 1.39%, while the male one is 1.27% (Anderson and Ray, 2010, Table 2). Since in high-income countries the male-female sex ratio is 1.25 in that age group (Anderson and Ray, 2010, Table 2), the number of missing females aged 1-5 in Sub-Saharan Africa is ((0.0139- (0.0127/1.25))*42 m. =) 0.16 million as shown in Table 2.

3. Low mortality reference populations and the sex ratio as a functional form of the reference standard

a. The problem

The flow measure of missing women employs male to female age-specific mortality rate ratios that prevail in high-income countries in 2000 as a reference standard. The main shortcoming of this reference standard is the very low overall mortality in high-income countries. In very low mortality environments, even the smallest differences in the numbers of deaths across sexes result in very high mortality sex ratios and therefore low female reference mortality rates and high flows of missing women. In the adult age group of 15 to 29 year-olds, which accounts for most of the estimated flow of missing women in Sub-Saharan Africa (see Table 2), the female mortality rate in high-income countries in 2000 is 0.42 per thousand whereas that for men is slightly higher at 1.08 per thousand, resulting in a sex ratio of 2.6; i.e. the probability of men dying in this young age group is 160% higher than the corresponding likelihood for women. By contrast, the female mortality rate ratio of 25 to 29 year old's in Sub-Saharan Africa is 14.8 and that of males is 10.8.

Table 2: Flow of missing women (in 1000s) by age group in 2000 using today's high-income countries' mortality rates as reference standard

	(1) Sub-Saharan Africa	(2) India	(3) China
At birth	0	184	644
0-1	32	146	109
1-4	160	164	23
5-14	70	93	2
15-29	578	258	24
30-44	345	93	73
45-59	84	120	89
60-79	213	541	490
80+	44	113	272
Total	1526	1712	1727
Share of females missing	0.47%	0.34%	0.31%

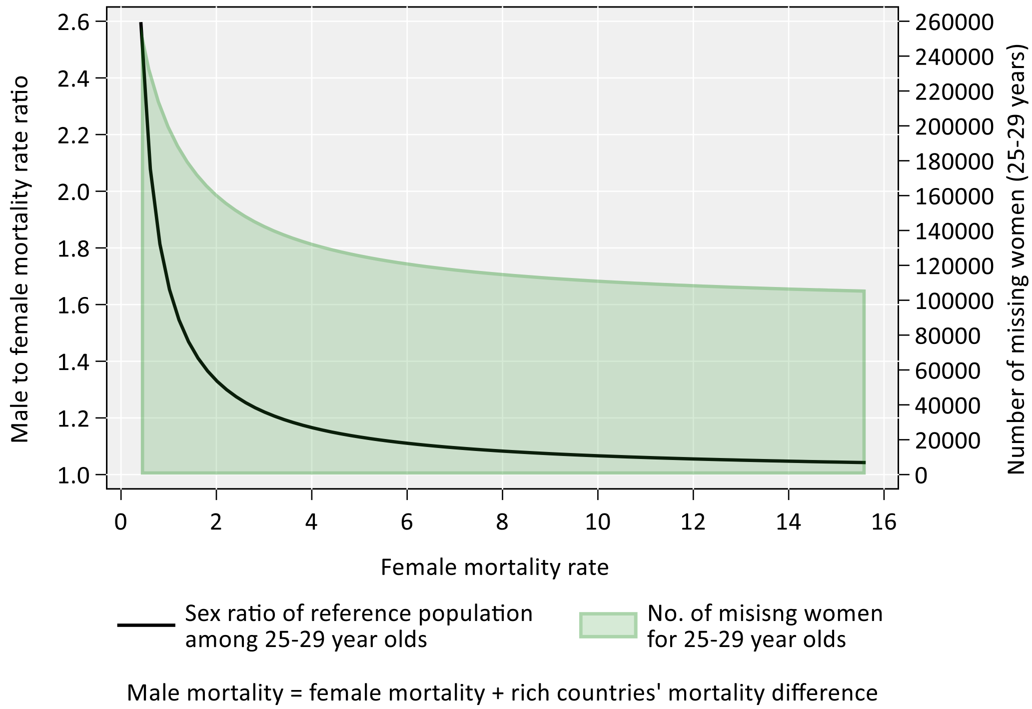
Note: Numbers do not sum to total because of rounding error.
Sources: Anderson and Ray (2010).

In Figure 1, we simulate how the mortality sex ratio of 25 to 29 year old's in the reference population would fall as the mortality level increases from very low mortality levels prevailing in high-income countries to high levels prevailing in Sub-Saharan Africa, while the difference in mortality rates between sexes remains constant. Figure 1 also shows the number of missing women associated with these mortality levels and sex ratios. Increasing the female mortality rate of 25 to 29 year old's only slightly from 0.42 to 2.42, reduces the sex ratio by half from 2.6 to 1.3 and the number of missing women falls from 258,118 to 152,434, a difference of over 100,000 just in the age group of 25 to 29 year old's alone. Further increases of the mortality level have much smaller impacts on the number of missing women as the mortality sex ratio is increasingly less sensitive as mortality levels rise. Increasing the female mortality rate of 25 to 29 year old's all the way to the level prevailing in Sub-Saharan Africa, which is 14.8, reduces the number of missing women to below 110,000.[6] This sensitivity of the number of missing women is particular to the flow measure because the stock measure does not rely on the number of deaths but on the number of males and females surviving which make up, of course, much larger numbers.

The above calculations illustrate the sensitivity of the number of missing women with respect to the mortality level for 25 to 29 year old's only, which is the age group with the highest mortality sex ratio. Figure 2 depicts the male to female mortality ratios for all age groups in high-income countries in 2000. The ratios are particularly high – above 2 – in the age groups between 15 and 35. Between the ages of 35 and 55 the mortality ratios are somewhat smaller averaging around 1.86. A final peak arises between the ages of 55 and 70 with mortality rate ratios over 1.94. For older ages the ratio falls sharply, but always remains above 1.

[6] At high mortality levels, also the difference between the male and female mortality rate is often larger. In Sub-Saharan Africa the difference is 3.9. If we were to use that difference instead, or increasing the difference from 0.66 to 3.9 as we increase the mortality level from 0.42 to 14.8, the implications of this simulation exercise wouldn't change.

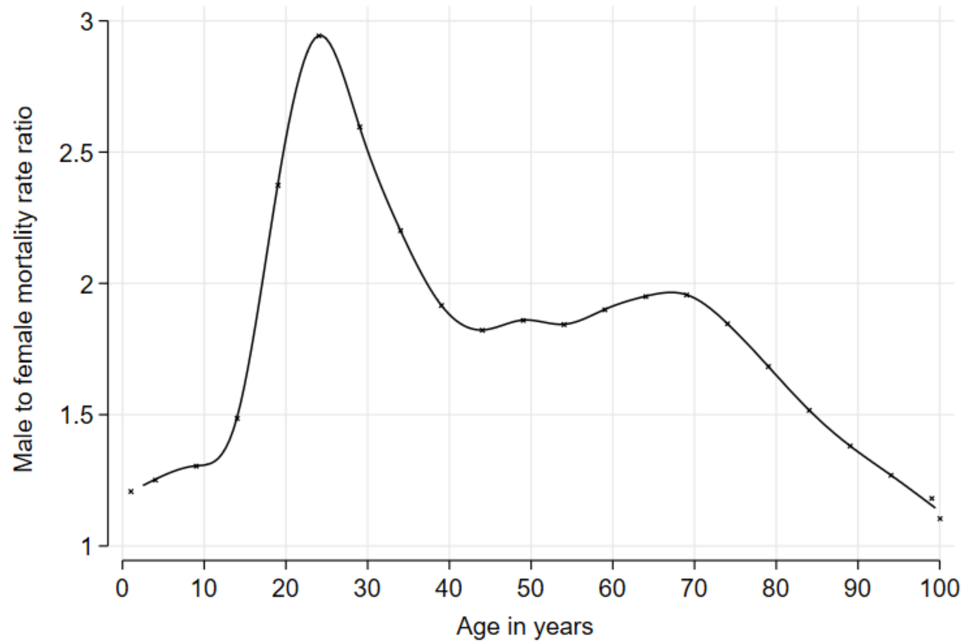
Figure 1: Female to male mortality rate ratio (sex ratio) of 25 to 29 year old's in reference populations and the associated number of missing women in the same age group



Note: Figure shows female to male mortality rate ratio (sex ratio) of 25 to 29 year old's in hypothetical reference populations with varying mortality levels, ranging from 0.42 (the female mortality rate in high-income countries) up to the female mortality rate in the same age group in SSA of 14.8 as well as the associated number of missing women in the age group of 25 to 29 year old's. The figure assumes a constant difference between the female and male mortality ratio equaling that of high-income countries. E.g., at a female mortality rate of 0.42, the male mortality rate is 1.08, and at a female mortality rate of 14.42, the male mortality rate is 15.08.

The first peak between 15 and 35 is largely a result of the role of injuries which particularly drives up male mortality rates in these young age groups; in the demographic literature this has sometimes been referred to as the \`testosterone spike' linked to injuries related to the combination of alcohol abuse, dangerous traffic behavior, and violence among young men (Kalben, 2000). Because of the overall low mortality among 15 to 29 year-olds in high-income countries, deaths from injuries are particularly prominent. In this young age group, injuries make up 71% of all deaths of males (and 51% of females) and the sex ratio of mortality from injuries is 3.7, whereas it is 1.6 for all other deaths. Therefore, the reference standard is heavily affected by this disproportionate importance of injuries in adult mortality for males in high-income countries.

Figure 2: Male to female relative mortality rate ratios by age in high-income countries in 2000



Source: Replicated from supplementary material of Anderson and Ray (2010)

The second peak is to a large extent caused by the mortality consequences of high rates of smoking among adults. As shown in detail by Crimmins, Preston and Cohen (2011), what matters for mortality are smoking rates two to four decades prior, i.e. smoking prevalence from 1960 to 1980 matters for mortality in 2000. As demonstrated by Deaton (2004) and Crimmins, Preston and Cohen (2011), smoking rates in high-income countries during those times were high, and much higher for males than for females. In fact, Crimmins, Preston and Cohen (2011) estimate that some 20% of all male deaths above age 50 in industrialized countries and less than 10% of all female deaths above age 50 were due to smoking-related diseases. As a result, men lost about 2.2 years of life expectancy due to smoking, compared to 1.2 for women. Many of these deaths (particularly related to lung cancer or cardiovascular diseases or pulmonary diseases) occur in the 55 to 70 age group, leading to the observed high sex ratio of mortality rates among 55 to 70 year-olds in high-income countries.[7]

[7] Since female smoking rates have been rising in high-income countries while overall smoking rates have fallen, the impact of gender gaps in smoking prevalence on sex-specific mortality rates will fall in coming decades. See Pampel (2003) and Crimmins, Preston and Cohen (2011) for a discussion. This also points to another problem of using the actual mortality rates in today's high-income countries as a standard. The number of missing women in Sub-Saharan Africa and India will automatically fall as a result of the greater gender balance in mortality due to smoking in high-income countries.

Of course, men in other parts of the world also tend to engage in riskier behaviors in comparison to women. However, the point is that these small differences in deaths from risky behavior across sexes cause very high sex ratios only in environments with very low overall mortality rates such as in high-income countries.

b. Use of different reference populations and alternative functional forms

To avoid the problems that come with low mortality level reference populations, one could either use a reference population with higher overall mortality rates or use a different functional form than the sex ratio for the reference standard.[8] As an alternative to the sex ratio, one could use the difference, i.e. subtracting the difference between the male and female mortality rates in the reference population, which is 0.66 (=1.08-0.42) for 25 to 29 year olds, from the actual male mortality rate to calculate the female reference mortality rate.[9] In this section, we test the sensitivity of the number of missing women using (a) reference mortality rates from Model Life Tables, referring back to the demographic literature for the estimation of missing women based on the stock measure (e.g. Coale, 1991; Klasen, 1994, Klasen, 1998, Klasen and Wink, 2002, 2003), and (a) the difference, rather than the ratio, in male and female mortality rates as a functional form.

Life Tables are tables that indicate for males and females separately and for each age the probability of dying before a person's next birthday and thus are a representation of death rates, i.e. the number of deaths at a specific age in a year divided by the number of survivors at that age and in that year. They are based on actual data for specific countries and time periods. Model Life Tables use these Life Tables from different datasets, countries and periods to estimate and provide Model Life Tables for environments with varying mortality patterns and conditions. There are various versions of Model Life Tables. The previous missing women literature mostly relied on the Princeton Model Life Tables (Coale, Demeny, and Vaughan, 1983), which are based on comprehensive and high-quality Life Tables from Europe, North America, and some non-European countries from about 1870

[8] Anderson and Ray (2010) briefly discuss functional forms, stating that they have a preference for the `scale neutrality embodied in the use of relative death rates' (Anderson and Ray, 2010: 1290).

[9] Note, we are not recommending the difference as a standard. The absolute difference would underestimate biological differences, particularly in infancy, and might overestimate these differences in older age groups where mortality levels are high. Also, the problem of male-biased injuries dominating the results would still be there (but muted through the functional form).

to 1960. These Life Tables were then combined to generate a set of Model Life Tables for different mortality conditions. The use of those Model Life Tables as a reference standard for generating estimates for missing women has several advantages (see Coale, Demeny, and Vaughan, 1983; Coale, 1991; Klasen, 1994; Klasen and Wink, 2002, 2003). First, they provide Life Tables for high-mortality environments such as Sub-Saharan Africa's but also medium or low mortality environments such as India's or China's.[10] Second, and closely related, they provide Life Tables for the entire spectrum of the epidemiological transition from predominantly communicable to predominantly non-communicable diseases. Third, they come from a time period where the main drivers of high male mortality today, especially related to nicotine abuse and traffic accidents, were much less serious and thus less likely to bias results. Lastly, they come from time periods for which the historical evidence suggests that gender discrimination in access to resources was comparatively low, certainly when compared to places such as India or China today (UN, 1998; Klasen, 1998; 2003; Johannsson, 1984, 1991; Schofield, 1991; Rosenzweig and Schultz, 1982; Klasen and Wink, 2003; Dyson and Moore, 1984; UN, 1998; Tabutin, 1978).[11] The technologies available to produce gender bias, particularly differential access to health care as well as sex-selective abortions, were unavailable or, in the case of health access, much less effective in producing gender differentials in mortality than today (e.g. Asfaw, et al. 2010; Klasen and Wink, 2003; Banister and Coale, 1994; Akbulut-Yuksel and Rosenblum, 2012).

Table 3 summarizes how the results on missing women by age change if one used either the difference in mortality rates in high-income countries (columns (2), (6) and (10)), the sex ratio of mortality rates from the Princeton Model Life Tables West (columns (3), (7) and (11)), or relied on the difference in mortality rates from the Princeton Model Life Tables

[10] They are less suitable for assessing very low mortality environments such as those prevailing in some industrialized countries today.

[11] There is a literature that has examined gender bias in mortality in historical Europe and that literature has indeed identified small episodes of such gender bias; but most episodes refer to the early 19th century, before the time period covered in the Model Life Tables; also the one episode in the late 19th century was focused on girls (aged 1-14) and has been linked to a particularly serious episode of tuberculosis that particularly affected them. But the presumed excess amounted to never more than 5-10% above expected mortality in this one age group and was very small in Scandinavia and the UK (and larger in Ireland). For a survey, see for example, Klasen (1998, 2003), Johansson (1984, 1991), among others.

West (columns (4), (8) and (12)).[12, 13] The findings based on the flow measure using the sex ratios of mortality in high-income countries today are reproduced and presented in columns (1), (5), and (9). The approach for identifying the number of missing women at birth remains the same throughout different reference standards (see Anderson and Ray, 2010).

Regardless of which alternative reference standard we use, the key findings are different than when using mortality sex ratios of high-income countries and they are different in similar ways across alternative reference standards. The first key finding, that the missing women problem is more severe in Sub-Saharan Africa than in India or China, reverses. The total and percentage flows of missing women are now much larger in India and China, compared to Sub-Saharan Africa. Using sex ratios as a functional form and Princeton Model Life Tables West as the reference population (column (3)), the number of missing women in Sub-Saharan Africa reduces from 0.47% of the female population or 1,526,000 women by a factor of one-seventh to 0.07% or 213,000 missing women. In India and China, the percentage share of missing women also reduces considerably, but by less, from 0.35% (1712) and 0.28% (1727), respectively, to 0.14% in both countries. Thus, the percentage share of missing women from the female population is half as large in Sub-Saharan Africa (0.07%) than in India and China (0.14%) using this alternative reference population, in comparison to about one-and-a half times as large using high-income countries as a reference population. Using the difference in mortality rates as a functional form leads to similar findings in Sub-Saharan Africa, India and China in terms of size of gender bias.[14]

[12] In a robustness check, we also use the Princeton Model Life Tables `North' with virtually identical results. They are available on request. Note that Anderson and Ray (2010) also report on results using model life tables in appendix B and also find that most missing women disappear in Sub-Saharan Africa but do not place any emphasis on these results in the text.

[13] We also used the sex ratio of mortality prevailing in Latin America as a robustness check and the results are available on request. Note that using Latin America does little to solve the problem as mortality rates in Latin America today are much lower than in Africa or India and in fact relatively close to high-income countries; the sex ratio of mortality in Latin America is even higher in some young adult age groups, related to particularly high male mortality rates due to injuries. So, the same concerns apply to this standard.

[14] Using the difference also shows qualitatively the same result; in fact, there are no more missing females in Sub-Saharan Africa. For reasons stated above, we consider the Princeton Model Life Tables and the expert standard to be the most reliable reference standard.

Table 3: Number of missing women (in 1000s) by age group using alternative reference standard

	Sub-Saharan Africa				India				China			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	RC Ratio	RC Diff.	MLT Ratio	MLT Diff.	RC Ratio	RC Diff.	MLT Ratio	MLT Diff.	RC Ratio	RC Diff.	MLT Ratio	MLT Diff.
At birth	0	0	0	0	184	184	184	184	644	644	644	644
0-1	32	-235	22	13	146	8	181	156	109	65	148	119
1-4	160	55	55	56	164	121	140	141	23	13	29	20
5-14	70	18	5	2	93	53	71	69	2	-9	12	4
15-29	578	279	241	237	258	127	111	113	24	-9	-9	-37
30-44	345	-56	-3	-51	93	-63	-70	-80	73	53	-2	-16
45-59	84	-144	-55	-93	120	-114	-39	-75	89	15	-9	-26
60-79	213	-18	-40	-40	541	103	76	41	490	170	27	12
80+	44	10	-10	-8	113	84	21	39	272	165	64	64
Total	1526	-91	213	116	1712	503	674	590	1727	1107	903	783
% of females	0.47	-0.03	0.07	0.04	0.35	0.10	0.14	0.12	0.28	0.18	0.14	0.13

Note: Ratio refers to the functional form of the reference standard that divides the male mortality rate in the region of interest by the sex ratio of mortality rates in the reference population. Diff. refers to the functional form of the reference standard that subtracts the difference in sex-specific reference population mortality rates from the male mortality rate in the region of interest. RC refers to high-income countries in 2000, MLT are Princeton Model Life Tables West.
Source: Anderson and Ray (2010) for population sizes and male and female deaths in Sub-Saharan Africa, India, China and high-income countries (referred to as EME) in 2000; Coale, Demeny, and Vaughan (1983) for Princeton Model Life Tables West.

The second key finding, that the problem of missing females in India is as much a problem among adults as it is among children, has also almost entirely vanished. The vast share of missing females in India emerges pre-birth and in the first few years of life. Using the Princeton Model Life Tables West as a reference population, the share of missing women accounted for by sex-selective abortions and elevated mortality rates in infancy and early childhood (<age 5) is over 70% compared to less than 30% when using high-income countries as a reference population (compare column (5) to columns (7) and (8)).

The third key finding, that the massive problem of gender bias in Sub-Saharan Africa is driven by excess mortality among adults, is also much reduced as the number of missing women reduces sharply in all age groups, changing the total from 1,526,000 to 213,000 missing women when Princeton Model Life Tables are used as a reference population instead of high-income countries. Although, it remains the case that a large number of women go missing in the 15 to 29 years age group. They make up 75% of all missing women in Sub-Saharan Africa (only taking into account positive numbers of missing women). In contrast, for the age groups of 30 and older the number of missing women even turns negative, suggesting little gender discrimination against middle aged and older adult women.

The distribution of missing women by age changes the least for China when the alternative reference standards are used. China already has a rather low mortality environment and thus the biases generated by using the ratios of mortality from high-income countries are less severe.

To summarize, using today's male to female mortality rate ratios in high-income countries likely overstates the flow of missing women in Africa and Asia. The total flow of missing women using the sex ratio in high-income countries as the reference standard is about 5 million which reduces to around 1.5-1.8 million per year when alternative reference standards are used. The estimates based on the alternative reference standards confirm previous estimates of missing women based on the stock measure; specifically, that the problem of missing women is more severe in India and China than in Sub-Saharan Africa and that it is largely driven by gender bias pre-birth and in young ages in India and China. Additionally, the alternative reference standards suggest that it is in particular adolescent and young adult women that have a mortality disadvantage rather than other adult age groups and children in Sub-Saharan Africa.

4. Controlling the reference standard for sex-specific differences in the disease environment

a. AIDS mortality

The disease environment in high-mortality regions such as Sub-Saharan Africa and, to a lesser extent, India is dramatically different from the disease environment in today's high-income countries. In Sub-Saharan Africa communicable diseases account for over 50% of all deaths.[15] On the contrary, in today's high-income countries, communicable diseases account for less than 5% of total deaths, whereas non-communicable diseases drive mortality in older ages and injuries in younger ages. But not only does the distribution of causes of death differ across countries, but also the distribution of deaths by sex across

[15] The discussion is based on numbers of death by disease, age and region reported in the web appendix and online supplementary material of Anderson and Ray (2010).

causes of death. For example, in high-income countries in 2000 the sex ratio of deaths from AIDS was on average 4 across adult age groups, i.e. men died at four times the rate of women from AIDS. In contrast, in Sub-Saharan Africa, the male to female ratio of deaths from AIDS was between 1 and 1.6 in most adult age groups, but only about 0.53 in the 15-29 age group. Thus, young adult women have higher mortality from AIDS than men. The stark difference in sex ratios of AIDS mortality across regions emerges because AIDS incidence and fatalities in high-income countries are mostly driven by gay men and IV drug users (also predominantly men), whereas heterosexual transmission plays a significant role in Sub-Saharan Africa (World Bank, 1997). Sex-specific differences in mortality across countries also exist, as previously discussed, for deaths from risky behaviors. However, whereas sex-specific differences in risky behavior are likely similar in Sub-Saharan African countries and high-income countries, the patterns for other diseases, such as AIDS evidently diverge to a large extent.

To account for differences in disease environments, the original flow measure neglects sex-specific differences and builds up the number of missing women disease group by disease group. Specifically, the sex ratio of age-specific mortality rates for each disease group in high-income countries is used as the reference standard, and the age- and disease-specific flows are then summed up over all ages and diseases to achieve the total number of missing women in a year. Building up the flow of missing women disease-by-disease, only has minor implications for the estimated total number of missing women. In Sub-Saharan Africa, the annual flow falls from 1.526 million to 1.385 million, in India from 1.712 to 1.637 million, and in China from 1.727 to 1.592 million (Anderson and Ray, 2010). The key findings of the overall high magnitude (4.6 million per year), the preponderance among adults, and the larger share of missing women in Africa, are not affected significantly by this disease correction.

The reference standard used in this disease correction presumes a counterfactual that without gender bias, AIDS in Sub-Saharan Africa would have also been primarily a disease affecting gay men and IV drug users and suggests that 58% of all adult missing women in Sub-Saharan Africa die from AIDS (see Table 4, column (3)). This counterfactual is likely false. AIDS in Africa has been a disease that has spread mostly through heterosexual intercourse (and mother to child transmission) due to the early emergence of the disease in Africa and its spread among heterosexual couples long before it was identified as a disease, together with a very high incidence of STDs (Iliffe, 2006; World Bank, 1997; Oster, 2005). It is likely to be true that the particularly high incidence and fatality of AIDS among young women is to some extent due to gender discrimination (e.g. unwanted sexual intercourse with older men), but this is not the main reason why AIDS is a disease affecting mostly heterosexual people in Sub-Saharan Africa across adult age groups. Assuming the ratio from high-income countries as a reference standard, likely overstates the relevance of AIDS in causing missing women considerably. To test the sensitivity in the number of missing women to alternative assumptions for an AIDS-specific reference standard in environments with a predominantly heterosexual transmission path, we assume a sex ratio of mortality in high-income countries that weights mortality rates by the share of hetero- and homosexual populations. For example, the sex ratio of HIV among heterosexual men (84.1%) and women (86.9%) as well as homosexual men and women (defined as those who are not heterosexual) in the US, as one of the established market economies, is around 1 (Lansky et al., 2015; Centers for Disease Control and Prevention, 2019).[16] This number is much smaller than the actual male to female mortality ratio from AIDS of 4 prevailing in most adult age groups of high-income countries in 2000. The sex ratio of AIDS mortality of 1 would still allow for high female mortality from AIDS in the 15-29 age group -- in which gender discrimination in the transmission of AIDS through forced sexual intercourse likely matters -- as the male to female ratio of deaths from AIDS among 15 to 29 year-olds is only 0.5 in Sub-Saharan Africa.[17]

Table 4 presents the flow of missing women from AIDS using a sex ratio of mortality from AIDS of 1 in columns (4) and (5) and a sex ratio of 1.2 as a more conservative reference standard for AIDS in columns (6) and (7). The flow of missing women from AIDS reduces by about 400,000 excess female deaths, from 611,000 to 19,000-171,000, essentially all

[16] Specifically, the sex ratio is calculated as $sex\ ratio = \frac{(share\ hetero.\ men * HIV\ cases\ per\ 100{,}000\ hetero.\ men) + (share\ homo.\ men * HIV\ cases\ per\ 100{,}000\ homo.\ men)}{(share\ hetero.\ women * HIV\ cases\ per\ 100{,}000\ hetero.\ women) + (share\ homo.\ women * HIV\ cases\ per\ 100{,}000\ hetero.\ women)}$.

[17] While partly this high mortality could indeed be due to discrimination linked to unwanted sexual intercourse of young women with older men, part of the high female mortality could also be due to earlier voluntary sexual activities of females (and thus earlier AIDS risks) which is observable in most countries of the world, including high-income countries; see World Bank (2012) for a discussion.

accruing in the 15-29 age group. Further, the share of adult women that go missing due to AIDS reduces from 58% to a more realistic percentage share of 4 to 28%.

Table 4: Number of missing women from AIDS (in 1000s) by age group using alternative reference standards

	Sub-Saharan Africa: Flow of missing women…						
	…total	**… from AIDS**					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Reference: Sex ratio, high-income countries Method: disease-by-disease group	Reference: Sex ratio, high-income countries		Reference: Sex ratio=1		Reference: Sex ratio=1.2	
Age	No.	No.	% of total	No.	% of total	No.	% of total
15-29	406	277	68%	180	58%	253	66%
30-44	278	240	86%	-91	--	-55	--
45-59	134	78	58%	-58	--	-112	--
60-79	217	15	7%	-11	--	6	--
80+	25	0	--	0	--	7	--
Total	1,060	611	58%	19	4%	171	28%

Note: Column (1) presents the total number of missing women when calculated disease group by disease group as presented in Table 6 of Anderson and Ray (2010). Columns (2) to (7) refer to different reference standards that were applied in the estimation of missing women from AIDS. Figures in columns (1), (2), (4) and (6) are rounded to the nearest thousand. Percentage shares in columns (3), (5) and (7) are based on numbers rounded to nearest thousand in columns (1), (2), (4) and (6).
Source: Anderson and Ray (2010) Table 6 for column (1) and Anderson and Ray (2010) supplementary material for population sizes and male and female deaths from AIDS in Sub-Saharan Africa and high-income countries (referred to as EME) in 2000.

b. Maternal mortality, injuries and suicide

There are numerous other diseases in addition to AIDS that require attention when the reference population is not drawn from a comparable disease environment. Some diseases are particularly rare in high-income countries, which renders the sex ratio of mortality prevailing in high-income countries an overall unsuitable reference standard for that particular disease.

As there is no such thing as paternal mortality that can be used to derive expected maternal mortality, the flow measure of missing women assumes that any maternal death above the rate prevailing in high-income countries is an excess death. Because maternal mortality is rare in high-income countries maternal mortality among 15 to 45 year-olds accounts for the second most excess female deaths in Sub-Saharan Africa (after AIDS, which would change if the alternative reference standards for AIDS were used) and for most excess female deaths in India. Hence, the assumption of all maternal deaths being excess female deaths is a determining factor for the total number of missing women and the distribution of missing women across age groups in Sub-Saharan Africa and India.

Maternal mortality is typically related to the poor overall health conditions and services prevailing in many developing countries. Figure 3 presents the relationship between child mortality rates of both sexes, which is an indicator of overall health conditions in a country, and the maternal mortality ratio. Child and maternal mortality show a close fit, indicating that maternal mortality goes hand-in-hand with poor overall health. If Sub-Saharan African countries were outliers in this relationship, then this might be due to gender inequality in health access that goes beyond poor overall health conditions. The dashed lines in Figure 3 present the predicted maternal mortality ratios based on a simple linear regression model of maternal mortality on child mortality and a dummy for Sub-Saharan Africa. The results suggest that on average about 11-16% of the maternal mortality in Sub Saharan Africa is not explained by poor health conditions and might be due to gender bias.[18] Thus, a small number of deaths from maternal mortality - around 22,000 to 32,000 rather than all 200,000 maternal deaths - might indeed be due to gender inequality.

Another noteworthy cause of death are injuries, specifically suicides and fires, for which differences in the distribution of deaths by sex potentially account for a substantial number of missing women. The sex ratio of deaths from suicides in India is relatively balanced and likely driven by economic hardship of families (especially in rural areas), whereas deaths from suicides in high-income countries is male biased and driven by psychiatric conditions (e.g. Aditjanyee, 1986; Hiroeh et al., 2001; Mortensen et al., 2000; Nordentoft et al., 1993; Qin et al., 2000). Similarly, death from fires is heavily male-biased in high-income countries (sex ratio of adult mortality (15-59) of 2.1), whereas in India fires afflict relatively more women (sex ratio of adult mortality (15-59) of 0.44). An estimated 106,000 out of 163,000 fire related deaths in India in 2001 accrued to women (Sanghavi et al., 2009; Bhalla et al., 2020). It was further estimated that more women died from fires than giving birth and at a rate 20 times more than in the rest of the world. These fire-related deaths are said to be due to the use of open fires, candles, and kerosene during domestic chores as well as self-immolation and domestic violence (Sanghavi et al., 2009). Thus, a

[18] 11-16% is the confidence interval of the mean of the Sub-Saharan Africa fixed effect (81 deaths) divided by the predicted maternal mortality ratio (mean=13.5%, CI=10.8-16.2%).

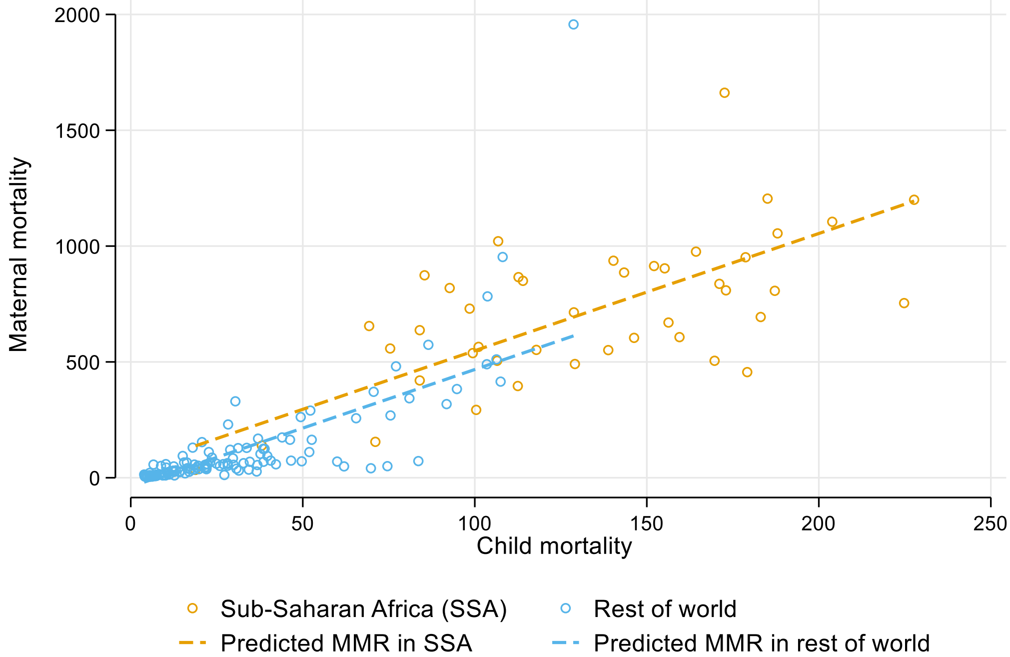
considerable share of fire related missing women is likely due to gender bias.[19] Yet, the flow of adult (>15 years) missing women from fires using the sex ratio from high-income countries as a reference standard is 72,000 in India, which make up over 50% of missing women from unintentional injuries and 80% of all fire related death of adult women would have to be due to gender discrimination, which seems unlikely.

The discussion illustrates that when using the sex ratio of mortality prevailing in today's high-income countries as a reference standard in the estimation of missing women, disease corrections must address the differences in the sex-specific disease environments between high-income countries and China, India, and Sub-Saharan Africa. Differences in disease environments may be based on different structures within disease groups (as in AIDS or injuries) or can be largely a result of the overall worse disease environment in developing countries (as is the case with maternal mortality).

Overall, the necessity for a detailed disease correction depends on the chosen reference population. The more appropriate the general reference population is, the less is such a disease correction required. For example, estimates using the Princeton Model Life Tables are much more similar to the within-disease corrected flow of missing women using high-income countries as a reference standard already because the Princeton Model Life Tables offer reference mortality rates for mortality environments that are similar to the country or region of interest.

[19] Due to death from interpersonal violence and women would also be less likely to conduct household chores if the division of domestic and wage work was equal across sexes.

Figure 3: Cross-country relationship between child mortality (x-axis) and maternal mortality (y-axis)



Notes: Child mortality is measured as the probability of dying between birth and exact age 5, expressed per 1,000 live births, while the maternal mortality ratio (MMR) is the number of maternal deaths per 100,000 live births. Predicted maternal mortality ratio is based on a linear prediction of maternal mortality on child mortality and a Sub-Saharan Africa dummy (both predictions result from the same model).
Source: UN Inter-agency Group for Child Mortality Estimation (UNICEF, WHO, World Bank, UN DESA Population Division; at child-mortality.org) for child mortality and Hogan et al. (2010) for maternal mortality ratio.

5. Availability and quality of mortality data from Sub-Saharan Africa

Putting the methodological question of appropriate reference standards aside, an important reason why flow estimates of missing women have not been produced before is the lack of reliable data on age-specific mortality rates by sex in many developing countries. Disease-specific data are even harder to come by as they require a complete vital registration system.

Our analysis follows previous estimations of the flow of missing women and uses mortality data for the year 2000 from the World Health Organization's 2002 Global Burden of Disease estimates (Anderson and Ray, 2010; World Bank,2012).[20] A detailed description of data sources and methods used in the derivation of the WHO mortality data are provided

[20] We directly use the age- and disease-specific mortality data of Anderson and Ray (2010) as provided in their supplementary materials.

in Mathers et al. (2004, see text, Table 4, Annex Table 6 and Annex Table 7). While the majority of countries have fairly recent data on infant and child mortality from DHS and related surveys (the modal time period being 1995-99, see Mathers et al. Table 3), data to estimate adult mortality are often lacking. In Sub-Saharan Africa, adult mortality estimates for 2000 are based on observations from only 4 out of 46 countries, two each for the early and the late 1990s (see Mathers et al., 2004, Table 4). In fact, reliable vital registration data are available for only one country in Sub-Saharan Africa, which is the island of Mauritius (see Mathers et al., 2004, Annex Table 6). Although China and India also lack full vital registration systems, in China age specific deaths could be estimated from a sample vital registration and the 2000 census, and in India data from a project vital registration system was available (see Mathers et al., 2004, pp.10/11 and Annex Table 6).

Since the Global Burden of Disease estimates in 2002, there have been multiple revisions and updates of the WHO life tables, the latest update being from December 2019. However, the availability of quality vital registration data has not changed since 2002.[21] Although a few more countries, for example Cabo Verde and South Africa, now collect and report vital registration data to the World Health Organization, they were excluded due to insufficient quality (World Health Organization, 2020 a,b). In consequence, even the most recent WHO mortality data for Sub-Saharan Africa only include high-quality vital registration data for Mauritius.

In the absence of vital registration data, age- and disease-specific mortality rates for Sub-Saharan Africa have to be estimated. This requires accurate calculations of overall age-specific mortality rates to provide the ‘envelope’ for total deaths that can then be attributed to different disease categories (Mathers et al., 2004). When data on child mortality (the likelihood of reaching age five) and adult mortality (the likelihood of reaching age 60 given that one has reached age 15) are available, modified logit life tables can be used to estimate mortality rates for all age groups (Murray et al., 2003 a, b). Murray et al. (2003a, b) demonstrate that the prediction method performs well in simulations if reliable data on child mortality and adult mortality for each sex are available; as the relationship between child and adult mortality differs between countries and across regions (see also Coale,

[21] See https://www.who.int/data/gho/data/themes/mortality-and-global-health-estimates/download-the-raw-data-files-of-the-who-mortality-database for “Availability Contains an Excel file with the list of countries-years available for the mortality and population data. Last updated: 15 December 2019.”

Demeny and Vaughan, 1983), these two pieces of information are crucial to generate the full set of age- and sex-specific mortality rates.

Because only child mortality data was available for the year of 2000 in Sub-Saharan Africa, the corresponding levels of adult mortality were essentially assumed, using historical model life tables, as explained by Mathers et al. (2004, p.11): “Based on the predicted level of child mortality in 2002, the most likely corresponding level of adult mortality (excluding HIV/AIDS deaths where necessary) was selected, along with uncertainty ranges, based on regression models of child versus adult mortality as observed in a set of almost 2000 life tables judged to be of good quality.” [22] There is no independent way of confirming these adult mortality rates generated by the life tables, let alone the relationship between male and female adult mortality rates.

In more recent years there has been some progress in the estimation of adult mortality rates (15-59). Since 2012, sex-specific adult mortality estimates of the Global Burden of Disease studies are based on sibling survival data in countries where no vital registration data are available, which in the case of Sub-Saharan Africa mostly come from Demographic and Health Surveys (Obermeyer et al., 2010; Wang et al., 2020, see supplementary appendix 1, section 2.3 and appendix table 3; World Health Organization, 2020 a,b). Albeit not perfect, these adult mortality estimates are based on actual data from Africa and therefore allow for a more reliable estimation of age-specific mortality rates using modified logit model life tables (Murray et al., 2003a,b; Wang et al., 2020; Wang et al. 2012, 2014; World Health Organization, 2020a,b).

Yet, the complete absence of vital registration systems makes it impossible to come up with precise age-specific mortality rates. In contrast, the stock estimates of missing women rely on much more widely available (and reliable) census information.

[22] Note that the 2000 life tables referred to here and lower down in the paper by Murray et al. (2003 a, b) are from all over the world, not from Africa alone.

6. Conclusions

The *flow measure* of missing women calculates excess female mortality that occurs in a certain time period, whereas the previous literature focused largely on calculating the *stock measure* of missing women, i.e. the deficit among alive women at one point in time. Estimations of excess female mortality by Anderson and Ray (2010) and World Bank (2012) suggest that gender bias in mortality is much larger than previously found (about 4 to 5 million excess female deaths per year vs. around 100 million missing women in total in 2000), is as severe among adults as it is among children in India, and is larger in Sub-Saharan Africa than in China and India. However, these estimates depend on particular methodological choices with respect to the reference standard and the method of disease correction. If alternative, potentially better suited, reference standards are used the results are aligned with those of the previous literature on the stock measure of missing women. The stock measure results show that more women are missing in China and India than in Sub-Saharan Africa and that, in China and India, missing women are largely attributable to excess female mortality before birth and/or the first few years of live. In addition to the question of appropriate reference standards, a lack of vital registration data in most developing countries makes it difficult to estimate precise numbers of excess female deaths.

Gender bias in mortality is a serious issue that deserves urgent attention by policy makers. According to our calculations, 1.5 to 1.8 million females perish in excess every year. Although, this number is much lower than the original flow estimates of 4 to 5 million yearly excess female deaths, our results suggest that the focus of the literature and associated policy proposals on South Asia and China, and on pre-birth and young children, is justified.

Our estimates confirm high excess female mortality in Sub-Saharan Africa among young adults, mainly driven by excess mortality from AIDS and maternal mortality, which provides some intriguing further detail to the issue of gender bias in mortality in Sub-Saharan Africa studied so far (e.g. Klasen, 1996a, b; Svedberg, 1996; Klasen and Wink, 2002; 2003; World Bank, 1997). Further investigation of gender bias in this age group in Sub-Saharan Africa should be a priority, as should be continuing efforts to improve vital data from developing countries to assess excess female mortality more accurately in the future.

References

Abrevaya, J. (2009). Are there missing girls in the United States? Evidence from Birth Data. *American Economic Journal: Applied Economics* 1 (2), 1-34.

Aditjanyee (1986). Suicide Attempts and Suicides in India: Cross-Cultural Aspects. *Int J Soc Psychiatry* 32, 64-73

Akbulut-Yuksel, M., and Rosenblum, D. (2012). The Indian Ultrasound Paradox. *IZA Discussion Paper*.

Anderson, S., and Ray, D. (2010). Missing Women: Age and Disease. *Review of Economic Studies* 77, 1262-1300.

Anderson, S., and Ray, D. (2012). The age distribution of missing women in India. *Economic and Political Weekly*, 87-95.

Anderson, S., and Ray, D. (2019). Missing Unmarried Women. *Journal of the European Economic Association* 17 (5), 1585-1616.

Asfaw, A., Klasen, S., and Lamanna, F. (2010). Intrahousehold health care financing strategy and the gender gap: Empirical evidence from India. *Health Economics* 19, 265-279

Banister, J., and Coale, A.J. (1994). Five Decades of Missing Females in China. *Demography* 31 (3), 459-479.

Bhalla, K., and Sanghavi, P. (2020). Fire-related deaths among women in India are underestimated. *The Lancet* 395 (10226), 779-780.

Bongaarts, J., and Guilmoto, C. Z. (2015). How many more missing women? Excess female mortality and prenatal sex selection, 1970–2050. *Population and Development Review* 41 (2), 241-269.

Case, A., and Paxon, C. (2005). Sex Differences in Morbidity and Mortality. *Demography* 42, 189-214.

Cawley J., and Ruhm, C. (2012). The Economics of Risky Health Behaviors. Chapter 3 in: Thomas G. McGuire, Mark V. Pauly, and Pedro Pita Barros (editors), *Handbook of Health Economics*, Volume 2, 95-199. Elsevier: New York.

Centers for Disease Control and Prevention. (2019). Estimated HIV incidence and prevalence in the United States, 2010–2016. *HIV Surveillance Supplemental Report 2019*; 24 (No. 1). http://www.cdc.gov/hiv/ library/reports/hiv-surveillance.html. Accessed [March 17, 2021].

Chen, G., and Oster, E. (2010). Hepatitis B does not explain male-biased sex ratios in China. *Economics Letters* 107 (2),142-144.

Coale, A.J. (1991). Excess female mortality and the balance of the sexes in the population: An estimate of the number of `missing females'. *Population and Development Review* 17 (3), 517-523.

Coale, A.J., Demeny, P., and Vaughan, B. (1983). Regional model life tables and stable populations. 2nd ed. New York: Academic Press.

Crimmins, E.M., Preston, S.H., and Cohen, B. (2011). Explaining Divergent Levels of Longevity in High-Income Countries". Washington, D.C.: The National Academies Press.

Das Gupta, M. (1987). Selective Discrimination against Females in Rural Punjab, India. *Population and Development Review* 13 (19), 77-100.

Das Gupta, M. (2005). Explaining Asia's "Missing Women": A New Look at the Data. *Population and Development Review* 31 (3), 529-535.

Das Gupta, M. (2006). Cultural versus biological factors in explaining Asia's Missing Women. *Population and Development Review* 32, 328-332.

Das Gupta, M., Chung, W., and Shuzhou, L. (2009). Is there an incipient turnaround in Asia's "missing girls" phenomenon?. *Policy Research Working Paper Series* 4846, The World Bank.

Deaton, A. (2004). Health in an Age of Globalization. In: Collins, S. and C. Graham (eds.) *Brookings Trade Forum*. Washington, D.C.: The Brookings Institution.

Dyson, T., and Moore, M. (1983). On kinship structure, female autonomy, and demographic behavior in India. *Population and Development Review* 9 (1), 35-60.

Ginsberg, C. (1989). Sex-specific mortality and the economic value of children in Nineteenth Centure Massachusetts. New York: Garland.

Hiroeh, U., Appleby, L., Mortensen, PB, and Dunn, G. (2001). Death by homicide, suicide, and other unnatural causes in people with mental illness: a population-based study. *Lancet* 358 (9299), 2110-2112.

Hogan, M.C, Foreman, K.J., Naghavi, M., Ahn, S.Y., Wang, M., Makela, S.M., Lopez, A.D., Lozano, R., and Murray, C.J.L. (2010). Maternal mortality for 181 countries, 1980–2008: a systematic analysis of progress towards Millennium Development Goal 5. *Lancet* 375, 1609-1623.

Iliffe, J. (2006). The African AIDS Epidemic: A History. Athens: Ohio University Press.

Johansson, S. (1984). Deferred Infanticide: Excess female mortality in childhood. In Hausfater, G., and Hrdy, S. (eds.) *Infanticide*. New York: Aladin, 1984, 463-486.

Johansson, S. (1991). Welfare, mortality and gender. Continuity and change in the explanation of male/female mortality differentials over 3 centuries. *Continuity and Change* 6, 135-177.

Johansson, S., and Nygren, O. (1991). The Missing Girls of China: A New Demographic Account. *Population and Development Review* 17 (1), 35–51.

Kalben, B.B. (2000). Why men die younger. *North American Actuarial Journal* 4 (4), 83-116.

Klasen, S. (1994). Missing Women Reconsidered. *World Development* 22, 1061-71 (1994).

Klasen, S. (1996a). Nutrition, Health, and Mortality in Sub Saharan Africa: Is There a Gender Bias?. *Journal of Development Studies* 32, 913–933.

Klasen, S. (1996b). Rejoinder. *Journal of Development Studies* 32, 944–948.

Klasen, S. (1998). Marriage, Bargaining, and Intrahousehold Resource Allocation: Excess Female Mortality Among Adults in Rural Germany (1740–1860). *Journal of Economic History* 58, 432–467.

Klasen, S. (1999). Family Composition, Gender preferences, and parental investment: Excess female mortality among infants and children during early German Development. Paper presented at the *ESF Conference 'Stayers and Leavers in Europe and Asia'*, The Hague, June 18-20, 1999.

Klasen, S. (2003). Weibliche Übersterblichkeit in Entwicklungsländern: Eine ökonomische Analyse. In: Ahrens, H. (ed.) *Neuere Ansätze der theoretischen und empirischen Entwicklungsforschung*. Berlin: Duncker&Humblot, 257-282.

Klasen, S. (2008). Economic Growth and Poverty Reduction: Measurement Issues using Income and Non-Income Indicators. *World Development* 36 (3), 420-445.

Klasen, S., and Wink, C. (2002). A Turning Point in Gender Bias in Mortality? an update on the number of missing women. *Population and Development Review* 28 (2), 285–312.

Klasen, S., and Wink, C. (2003). 'Missing Women': Revisiting the Debate. *Feminist Economics* 9 (23), 263–299.

Lansky, A., Johnson, C., Oraka, E., Sionean, C., Joyce, M. P., DiNenno, E., and Crepaz, N. (2015). Estimating the number of heterosexual persons in the United States to calculate national rates of HIV infection. *PloS one* 10 (7), e0133543.

Lin, M.-J., and Luoh, M.-C. (2008). Can hepatitis B mothers account for the number of missing women? Evidence from 3 million newborns in Taiwan. *American Economic Review* 98 (5), 2259-73.

Lopez, A.D., Ahmad, O.B., Guillot, M., Inoue, M., Ferguson, B.D., and Salomon, J.A. (2003). Life tables for 191 countries for 2000: data, methods, results. *GPE Discussion Paper* No. 40. Geneva: WHO.

Mathers, C., Bernard, C., Moesgaard Iburg, K., Inoue, M., Ma Fat, D., Shibuya, K., Stein, C.,Tomijima, N., and Xu, H. (2004). Global Burden of Disease in 2002: Data Sources, Methods, and Results. *Global Programme on Evidence for Health Policy Discussion Paper* No. 54. World Health Organization.

Mortensen, PB, Agerbo, E., Erikson, T., Qin, P., and Westergaard-Nielsen, N. (2000). Psychiatric illness and risk factors for suicide in Denmark. *Lancet* 355 (9197), 9-12.

Murray, C.J.L., Ferguson, B.D., Lopez, A.D., Guillot, M., Salomon, J.A., and Ahmad, O. (2003a). Modified logit life table system: principles, empirical validation, and application. *Population Studies* 57 (2), 165–182.

Murray, C.J.L., Ferguson, B.D., Lopez, A.D., Guillot, M., Salomon J.A., and Ahmad, O. (2003b). Modified Logit Life Table System: Principles, Empirical Validation and Application. *GPE Discussion Paper* No.39. Geneva: WHO

Murthi, M., Guio, A.-C., and Dreze, J. (1995). Mortality, Fertility, and Gender Bias in India: A District-Level Analysis. *Population and Development Review* 21 (4), 745-782.

Nordentoft, M., Breum, L., Munck, L.K., Nordestgaard, A.G., Hunding, A., and Laursen, P.A. (1993). High mortality by natural and unnatural causes: a 10 year follow up study of patients admitted to a poisoning treatment centre after suicide attempts. *British Medical Journal* 306, 1637.

Obermeyer, Z., Rajaratnam, J.K., Park, C.H., Gakidou, E., Hogan, M.C., Lopez, A.D., and Murray, C.J. (2010). Measuring adult mortality using sibling survival: a new analytical method and new results for 44 countries. 1974–2006. *PLoS Med* 7 (4), e1000260.

Oster, E. (2005). Sexually Transmitted Infections, Sexual Behavior, and the HIV/AIDS Epidemic. *Quarterly Journal of Economics* 120 (2), 467-515.

Pampel, F.C. (2003). Declining sex differences in mortality from lung cancer in high-income nations. *Demography* 40 (1), 45-65.

Qin, P., Mortensen, P.B., Agerbo, E., Westergard-Nielsen, N., and Eriksson, T. (2000). Gender differences in risk factors for suicide in Denmark. *British Journal of Psychiatry* 177, 546-550.

Rajaratnam, J.K., Marcus, J.R., Levin-Rector, A., et al. (2010). Worldwide mortality in men and women aged 15–59 years from 1970 to 2010: a systematic analysis. *Lancet* 375, 1704-1720

Rogers, R.G., Everett, B.G., Saint Onge, J.M., and Krueger, P.M. (2010). Social, Behavioral, and Biological Factors, and Sex Differences in Mortality. *Demography* 47, 555-578.

Rosenzweig, M.R., and Schultz, T.P. (1982). Market Opportunities, Genetic Endowments and Intrafamily Resource Distribution: Child Survival in Rural India. *The American Economic Review* 72 (4), 803-815

Sanghavi, P., Bhalla, K., and Das, V. (2009). Fire-related deaths in India in 2001: a retrospective analysis of data. *The Lancet* 373 (9671), 1282-1288.

Schofield, R. (1991). The decline of mortality in Europe. Oxford: Clarendon Press.

Sen. A. (1989). Women's Survival as a Development Problem. *Bulletin of the American Academy of Arts and Sciences* 43 (2), 14-29.

Sen, A. (1990). More Than 100 Million Women Are Missing. *New York Review of Books* 37.

Svedberg, P. (1996). Gender Bias in Sub-Saharan Africa: Reply and Further Evidence. *Journal of Development Studies* 32 (4), 933-943.

Tabutin, D. (1978). La surmortalité féminine en Europe avant 1940. *Population* (French Edition) 33 (1), 121-148

United Nations. (1998). Too young to die: genes or gender. New York: UN.

UN World Population Prospect Revision. (2000). https://population.un.org/wpp/Download/Archive/Standard/, downloaded April 3, 2021.

Waldron, I. (1976). Why Do Women Live Longer Than Men?. *Social Science and Medicine* 10, 349-362.

Waldron, I. (1982). Sex Differences in Human Mortality: The Role of Genetic Factors. *Social Science and Medicine* 17, 321-333.

Waldron, I. (1983). The Role of Genetic and Biological Factors in Sex Differences in Mortality. In: Lopez, A.D., and Ruzicka, L.T. (eds.) *Sex Differentials in Mortality*. 141-164. Canberra: Australian National University Press.

Waldron I. (1993). Recent Trends in Sex Mortality Ratios for Adults in Developed Countries. *Social Science Medicine* 36, 451-462.

Waldron, I. (1998). Sex Differences in Infant and Early Childhood Mortality. In: United Nations (ed.) *Too Young to Die – Genes or Gender?*. 64-83. New York: United Nations.

Wang H, Dwyer-Lindgren, L., Lofgren, K., et al. (2012). Age-specific and sex-specific mortality in 187 countries, 1970–2010: a systematic analysis for the Global Burden of Disease Study 2010. *The Lancet* 380 (9859), 2071–94.

Wang, H., Liddell, C., Coates, M., et al. (2014). Global, regional, and national levels of neonatal, infant, and under-5 mortality during 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *The Lancet* 384 (9947), 957–79.

Wang, H., Abbas, K.M., Abbasifard, M., Abbasi-Kangevari, M., Abbastabar, H., Abd-Allah, F., ... and Damiani, G. (2020). Global age-sex-specific fertility, mortality, healthy life expectancy (HALE), and population estimates in 204 countries and territories, 1950–2019: a comprehensive demographic analysis for the Global Burden of Disease Study 2019. *The Lancet* 396 (10258), 1160-1203.

World Bank. (1997). Confronting AIDS. Washington DC: The World Bank.

World Bank. (2012). World Development Report 2012: Gender Equality and Development. Washington DC: The World Bank.

World Health Organization. (2000a). WHO methods and data sources for life tables 1990-2019. *Global Health Estimates Technical Paper* WHO/DDI/DNA/GHE/2020.1. Geneva: World Health organization.

World Health Organization (2000b). WHO methods and data sources for country-level causes of death 2000-2019. *Global Health Estimates Technical Paper* WHO/DDI/DNA/GHE/2020.2. Geneva: World Health Organization.

World Health Organization. (2001). The World Health Report 2001: Mental health: new understanding, new hope. Geneva: World Health Organization.

World Health Report (WHR). (2001). Mental Health: New Understanding, New Hope. World Health Organization.